

**СОЗДАНИЕ ВИРТУАЛЬНОЙ КЛИЕНТСКОЙ БАЗЫ ДЛЯ АНАЛИЗА  
КРЕДИТОСПОСОБНОСТИ РОССИЙСКИХ ПРЕДПРИЯТИЙ**  
**Creating of virtual clients data base for solvency analysis of Russian enterprises**

**И.В. ШЕВЧЕНКО,**

*доктор экономических наук, профессор, декан экономического факультета, завкафедрой  
мировой экономики*

**А.А. ХАЛАФЯН**

*кандидат технических наук, доцент, профессор кафедры прикладной математики*

**Е.Ю. ВАСИЛЬЕВА**

*аспирант кафедры прикладной математики  
Кубанский государственный университет*

*Shevchenko I.V., professor,*

*Halaphyan A.A., associate professor,*

*Vasileva E.Y., graduate student,*

*Kuban State University*

**Аннотация**

В настоящее время актуальна проблема определения риска банкротства предприятия. Решение проблемы возможно путем использования статистических классификационных и эвристических методов, которые требуют наличия обучающей выборки – совокупности данных по предприятиям, уровень риска банкротства которых изначально известен. На практике сбор такой информации является весьма трудоемким и доступ к ней либо ограничен, либо вообще отсутствует. В статье предложен метод создания виртуальной клиентской базы, который основан на выявленных закономерностях значений финансовых показателей предприятия. При помощи клиентской базы посредством таких методов как дискриминантный анализ, деревья классификации, нейронные сети, можно оценить величину риска банкротства любого предприятия, если известны определенные показатели его экономической деятельности.

**Abstract**

Now days there are an urgency problem of enterprises non-payment risk. It is possible to solve a problem by using classification statistic methods, that based on instruction sample – total

combination of enterprises data, that credit risk levels are known. It's highly difficult to gather such information cause of limited data access or banning access at all. In the present work the creating of virtual clients data base method is suggested. This one is based on financial indices patterns, that are known. The risks of insolvency can be appreciated with virtual clients data base with using discriminant analysis method, classification tree models, neural networks methods.

**Ключевые слова:** классификационный анализ, обучающая выборка, кластерный анализ, дискриминантный анализ, нейронные сети, лингвистическая шкала, рентабельность, платежеспособность, теория нечетких множеств

**Keywords:** classification analysis, instraction sample, cluster analysis, discriminant analysis, neural networks methods, linguistic scale, profitability, solvency, fuzzy sets approach.

В настоящее время, остро встала проблема информационной неопределенности предприятий, как объекта исследований. Наиболее существенной данная проблема является для кредитных организаций при определении уровня риска клиента в процессе осуществления финансового мониторинга.

Вопрос кредитоспособности предприятия на данный момент является актуальным. Существует множество различных методов и подходов для определения уровня риска банкротства. Но далеко не все эти методы являются приемлемыми из-за недостатка исходных данных и неустойчивости финансовых показателей предприятия (параметров). Очевидно, что значения финансовых показателей не только изменяемы во времени, но и зависимы от территориального фактора самого предприятия. Объясняется это тем, что условия развития предприятий в различных регионах страны различаются. Например, широко известный метод Альтмана определения уровня риска предприятия, не является устойчивым к вариациям исходных данных, так как эти данные обладают сильной динамикой изменения как во времени, так и в пространстве.

Наиболее удачным способом решения проблемы выбора устойчивых во времени и в пространстве финансовых показателей является метод комплексного финансового анализа предприятий, предложенный в статье [1]. Авторами был получен ряд самых общих закономерностей, которым подчиняются финансовые показатели российских предприятий. Это связано с тем, что комплексные финансовые оценки были выбраны авторами работы из поля качественных, а не количественных признаков. В основе метода, представленного в [1], лежит понятие нечетких множеств, что позволяет установить

логическую связь количественных значений показателей с некоторым комплексным показателем, характеризующим финансовое состояние в целом. Взаимосвязь комплексного показателя и количественных признаков проанализирована в статье [2]. В работе [3] впервые были предложены показатели, наилучшим образом характеризующие состояние предприятий в условиях российской экономики. Они отражают мнения руководителей коммерческих предприятий (по результатам их опроса) или входят в состав методики Федерального управления по делам несостоятельности (банкротства) и моделей Альтмана. После анализа собранных авторами [3] материалов ими были отобраны 16 показателей. Позже, уже дополненные и расширенные, они легли в основу метода оценки финансового состояния предприятий, описанного в статье [1].

Целью настоящей работы является генерация виртуальной клиентской базы предприятий на основе метода оценки финансового состояния, построенного на нечетких множествах. База предприятий может служить обучающей выборкой для методов классификационного анализа при определении уровня риска реально существующего предприятия. Имея клиентскую базу, в которой присутствуют виртуальные предприятия, принадлежащие к различным группам риска, можно такими методами как, нейронные сети, деревья классификации, или дискриминантный анализ, прогнозировать принадлежность предприятия к той или иной группе риска. Так же, далее в статье будет показано, что созданная база, как совокупность предприятий, может быть корректно разбита на уровни выбранной нами лингвистической шкалы параметров предприятия.

При рассмотрении параметров предприятий необходимо учесть, что используемый нами метод для создания виртуальной клиентской базы применим в случае макроэкономической стабилизации, то есть периода когда наблюдается стабильный рост валового внутреннего продукта (ВВП) одновременно с низкими уровнями колебания доллара-евро и процента инфляции. Произведя анализ данных показателей, представленных на сайте Международного валютного фонда [4], был выделен период с 2000 г.(высокий показатель инфляции) по III квартал 2008 г. включительно (мировой финансовый кризис). Данный период дает возможность говорить об однородной финансовой статистики. Здесь однородность подразумевается, как наличие незначительных изменений, а не как равенство показателей.

Для создания клиентской базы мы воспользовались 16 наиболее весомыми экономическими показателями, образующими четыре группы [1].

**Группа 1** – рентабельность:

**R1** – общая рентабельность (балансовая прибыль/выручка от продаж + внереализационные доходы, % );

**R2** – рентабельность активов (чистая прибыль/средняя балансовая стоимость активов, % за кв);

**R3** – рентабельность собственного капитала (чистая прибыль/ капитал и резервы (за вычетом собственных акций, выкупленных у акционеров) – целевое финансирование и поступления + доходы будущих периодов, % за кв);

**R4** – рентабельность продукции (вычисляется как «прибыль от продаж/выручка от продаж, %);

**R5** – рентабельность оборотных активов (чистая прибыль/средняя стоимость активов, % за кв).

**Группа 2** – ликвидность и платежеспособность:

**L1** – быстрый коэффициент ликвидности (оборотные активы – запасы – налог на добавленную стоимость по приобретенным ценностям – долгосрочная дебиторская задолженность)/краткосрочные обязательства(не включая доходы будущих периодов), безразмерная величина);

**L2** – коэффициент покрытия запасов (оборотные собственные средства + краткосрочные займы и кредиты + краткосрочная кредиторская задолженность)/средняя величина запасов, %);

**P1** – текущий коэффициент ликвидности (оборотные активы – долгосрочная дебиторская задолженность)/краткосрочные обязательства (не включая доходы будущих периодов), безразмерная величина).

**Группа 3** – деловая активность:

**A2** – оборачиваемость активов (выручка от продажи (за вычетом налога на добавленную стоимость, акцизов и др. обязательств)/ средняя стоимость активов, раз за квартал);

**A4** – оборачиваемость кредиторской задолженности (выручка от продажи без учета коммерческих и управленческих расходов/средняя кредиторская задолженность, раз за квартал);

**A5** – оборачиваемость дебиторской задолженности (выручка от продажи (за вычетом налога на добавленную стоимость, акцизов и др. обязательств)/(дебиторская задолженность на конец отчетного периода – задолженность учредителей по вкладам в уставной капитал на конец отчетного периода), раз за квартал);

**A6** – оборачиваемость запасов (себестоимость/средняя величина запасов, раз за квартал).

**Группа 4** – финансовая устойчивость:

**F1** – коэффициент финансовой зависимости (долгосрочные обязательства + краткосрочные обязательства(не включая доходы будущих периодов)/(капитал и резервы(за вычетом собственных акций, выкупленных у акционеров) – целевое финансирование и поступления + доходы будущих периодов, безразмерная величина);

**F2** – коэффициент автономии собственных средств (капитал и резервы (за вычетом собственных акций, выкупленных у акционеров) – целевое финансирование и поступления + доходы будущих периодов/внеоборотные активы + оборотные активы, безразмерная величина);

**F3** – обеспеченность запасов собственными оборотными средствами (собственные оборотные средства/запасы, безразмерная величина);

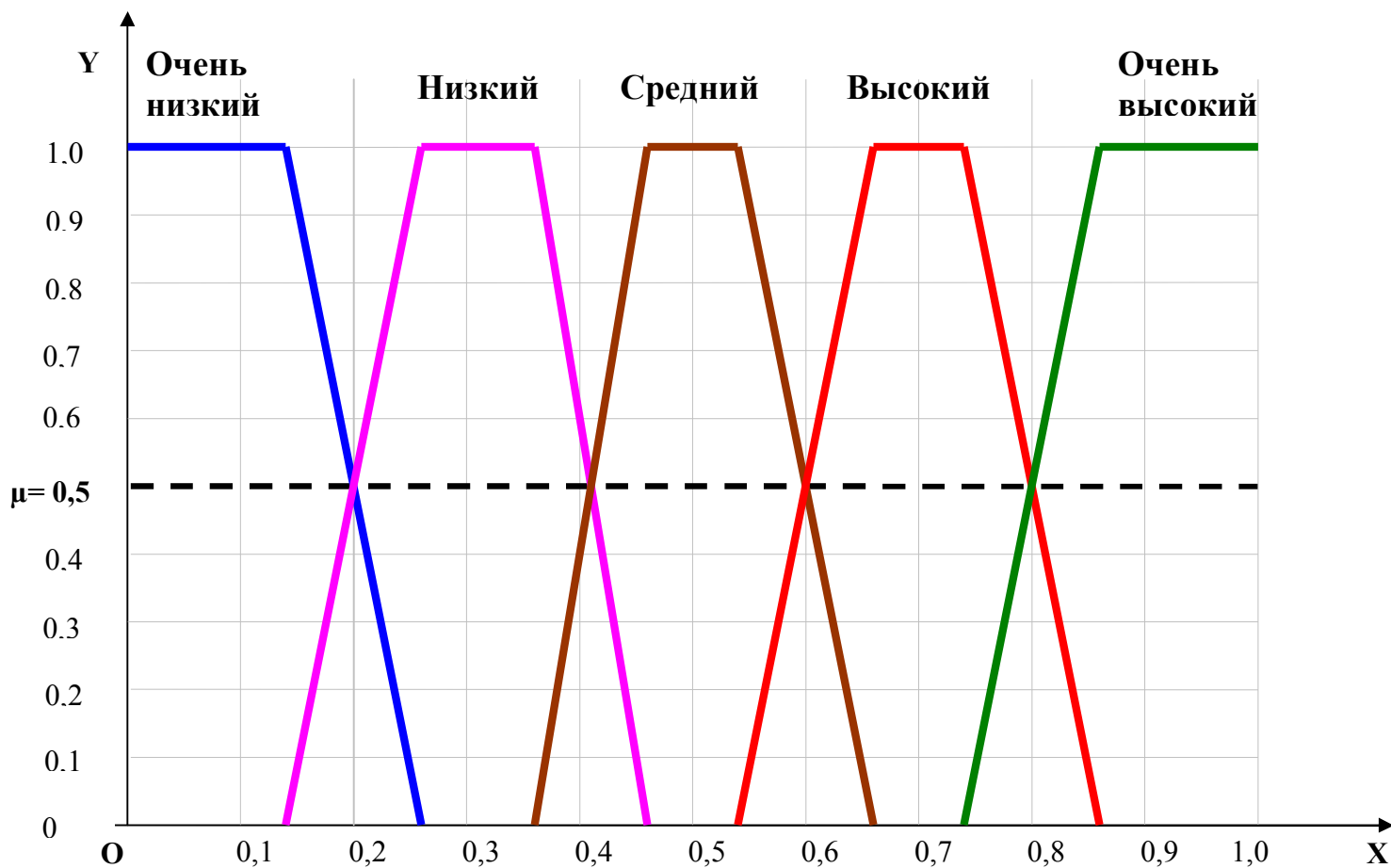
**F4** – индекс постоянного актива (внеоборотные активы + долгосрочная дебиторская задолженность/капитал и резервы (за вычетом собственных акций, выкупленных у акционеров) – целевое финансирование и поступления + доходы будущих периодов, безразмерная величина).

Для качественной и более наглядной интерпретации уровней выбранных показателей воспользовались лингвистической шкалой [5] – пенташкалой, т.е. пятиуровневым классификатором, где уровни интервальных значений параметров представлены в пяти вариациях: «Очень низкий (ОН) – Низкий (Н) – Средний (Ср) – Высокий (В) – Очень высокий (ОВ)». Лингвистическая шкала является наиболее наглядной для принятия решения специалистом. Наряду с пятиуровневым классификатором также используют бинарные («Низкий – Высокий») и тринарные («Низкий – Средний – Высокий») шкалы. Преимущество пенташкалы в нашем случае заключается в том, что она является более развернутой, чем бинарные и тринарные шкалы и в то же время четко определяет степень выбора того или иного решения. Шкалы состоящие из 7 элементов сильно рассеивают внимание и являются значительно более громоздкими. Именно поэтому, при определении уровня риска предприятия наиболее оптимальным является пятиуровневый классификатор.

Для пенташкалы была построена система функций, которые соответствуют нечетким подмножествам. Простейшим способом задания является система трапециевидных нечетких чисел. На рисунке 1 приведен пример такой системы с нормированным значением показателя (т.е. диапазон изменения значения финансового показателя находится в пределах от 0 до 1).

Чем ниже уровень показателя (рисунок 1), тем выше уровень риска банкротства предприятия. Трапеции классификатора, являющиеся функциями пенташкалы, попарно пересекаются в точках, у которых совпадают ординаты (0,5). Ось абсцисс отражает

величину значения параметра, а ось ординат – уверенность специалиста в принятии решения. В точках пересечения прямой  $\mu = 0,5$  и трапеций достигается максимум информационной неопределенности, когда соседние качественные свойства обладают одинаковой различимостью. Для остальных точек лежащих на оси ординат можем наблюдать нарастание силы одного из качественных признаков и одновременное убывание (в том же темпе) смежного качественного признака. Данная шкала, как шкала Поспелова, является непротиворечивым и полным классификатором [5], а значит и лингвистический анализ на ее основе непротиворечив.



Ось **OX** – нормированное значение экономического показателя  
 Ось **OY** - вероятность отнесения специалистом исследуемое предприятие к соответствующей группе уровня риска.

Рисунок 1. Пример системы пенташкалы с нормированным значением экономического показателя.

При помощи рассмотренной классификации было выделено пять групп:

1 группа – предприятия очень высокого уровня риска банкротства, которым соответствуют очень низкие уровни классификатора каждого из 16 рассматриваемых параметров;

2 группа – предприятия, находящиеся в зоне высокого риска банкротства, которым соответствует низкий уровень значений классификатора параметров;

3 группа – предприятия, находящиеся в зоне среднего риска банкротства, соответствующие среднему уровню значений классификатора параметров;

4 группа – предприятия, находящиеся в зоне низкого риска банкротства, соответствующие высокому уровню значений классификатора параметров;

5 группа – кредитоспособные предприятия, которым можно выдать кредит, соответствующие очень высокому уровню значений классификатора параметров.

Используя пенташкалу для каждого из показателей, составили таблицу финансовых показателей, четко определяющих принадлежность предприятия к той или иной группе. Так как ось ординат отражает уверенность специалиста при отнесении параметра к одному из пяти уровней, то для формирования четких классов приписывали значение параметра к тому уровню, у которого ордината соответствующей трапеции для данного значения больше. Если ординаты совпадали, то значение параметра относили к «более низкому» лингвистическому уровню из них.

Таблица 1 отражает диапазоны значений финансовых показателей, соответствующие каждому из уровней кредитоспособности предприятия.

Таблица 1. Диапазон значений экономических показателей по группам уровня риска

Группы уровня риска	Экономические показатели							
	L1	L3	P1	F1	F2	F3	F4	R1
Размерность показателя	Нет	%	нет	нет	нет	нет	нет	%
Очень высокий риск	0,1 – 0,2	40 – 77	0,1 – 0,4	2,70 – 2,50	0,25 – 0,31	-2,00 – -1,66	1,60 – 1,51	0 – 1
Высокий риск	0,2 – 0,6	77 – 144	0,4 – 0,8	2,50 – 1,70	0,31 – 0,54	-1,66 – -0,33	1,51 – 1,18	1 – 5
Средний риск	0,6 – 1,2	144 – 267	0,8 – 1,5	1,70 – 1,17	0,54 – 0,83	-0,33 – 0,92	1,18 – 0,83	5 – 16
Низкий риск	1,2 – 2,0	267 – 454	1,5 – 2,4	1,17 – 0,90	0,83 – 0,93	0,92 – 1,84	0,83 – 0,56	16 – 27
Очень низкий риск	2,0 – 2,2	454 – 580	2,4 – 2,8	0,90 – 0,50	0,93 – 0,95	1,84 – 3,50	0,56 – 0,20	27 – 30
Группы уровня риска	Экономические показатели							
	R2	R3	R4	R5	A2	A4	A5	A6
Размерность показателя	% за кв.	% за кв.	%	% за кв.	раз в кв.	раз в кв.	раз в кв.	раз в кв.
Очень высокий риск	-3,00 – -2,50	-4,00 – -3,33	0 – 1	-5,00 – 4,42	0,06 – 0,08	0,40 – 0,50	0,60 – 0,70	1,0 – 1,3
Высокий риск	-2,50 – 0,50	-3,33 – 0,66	1 – 5	-4,42 – 0,58	0,08 – 0,15	0,50 – 0,90	0,70 – 1,10	1,3 – 2,3
Средний риск	-0,50 – 1,57	-0,66 – 3,17	5 – 15	-0,58 – 2,58	0,15 – 0,24	0,90 – 1,54	1,10 – 1,74	2,3 – 3,5

Низкий риск	1,57 – 4,34	3,17 – 7,84	15 – 27	2,58 – 10,27	0,24 – 0,29	1,54 – 2,07	1,74 – 2,27	3,5 – 5,5
Очень низкий риск	4,34 – 8,00	7,84 – 20,00	27 – 46	10,27 – 18,00	0,29 – 0,58	2,07 – 5,80	2,27 – 5,80	5,5 – 15

Из 16 представленных показателей значения F1 и F4 являются *инверсными* по отношению ко всем остальным показателям, так как с ростом количественных уровней этих параметров финансовое положение предприятия ухудшается

При помощи средств программы Statistica 6 функцией генерации нормально распределенных случайных чисел были занесены в базу значения для 1000 предприятий пяти классов (уровней риска) по всем показателям. При введении аргументов данной функции генерации в качестве математического ожидания были выбраны середины интервалов, указанных в таблице, а в качестве среднеквадратического отклонения – разность между верхней границей интервала и средним значением деленная на три. При создании базы искусственно генерировали по 200 предприятий каждого класса.

Чтобы убедиться в том, что полученные данные определяют принадлежность к обозначенным классам, провели кластерный анализ, смысл которого заключается в разбиении множества исследуемых объектов и признаков на однородные в некотором смысле группы, или кластеры.

Наша задача заключалась в проверке схожести предприятий, принадлежащих одному классу и максимальной различимости предприятий различных классов. Для этого был проведен кластерный анализ с различным заданием количества кластеров. При разбиении был учтен критерий оптимальности: каждое предприятие должно принадлежать одному и только одному подмножеству разбиения (кластеру); предприятия, принадлежащие одному и тому же кластеру, должны быть сходными; предприятия, принадлежащие разным кластерам, должны быть разнородными.

Этот критерий представляет собой *целевую функцию* – функционал, выражающий уровни желательности различных группировок. Сходство и разнородность объектов определяется при помощи  $m$ -мерного евклидова расстояние между векторами измерений.

Для анализа использовали метод  $k$ -средних (*k-means clustering*), предполагая, что число максимально различных групп/кластеров  $m$  по всей совокупности 1000 наблюдений/предприятий. В качестве значения  $m$  взяли число 5. Результат анализа показал, что все предприятия попали в кластеры, каждый из которых соответствует определенному классу уровня риска, число которых так же равно 5. На рисунке 2 изображены графики средних значений всех 16 показателей для каждого из 5 кластеров. Каждому кластеру соответствует график определенного цвета.



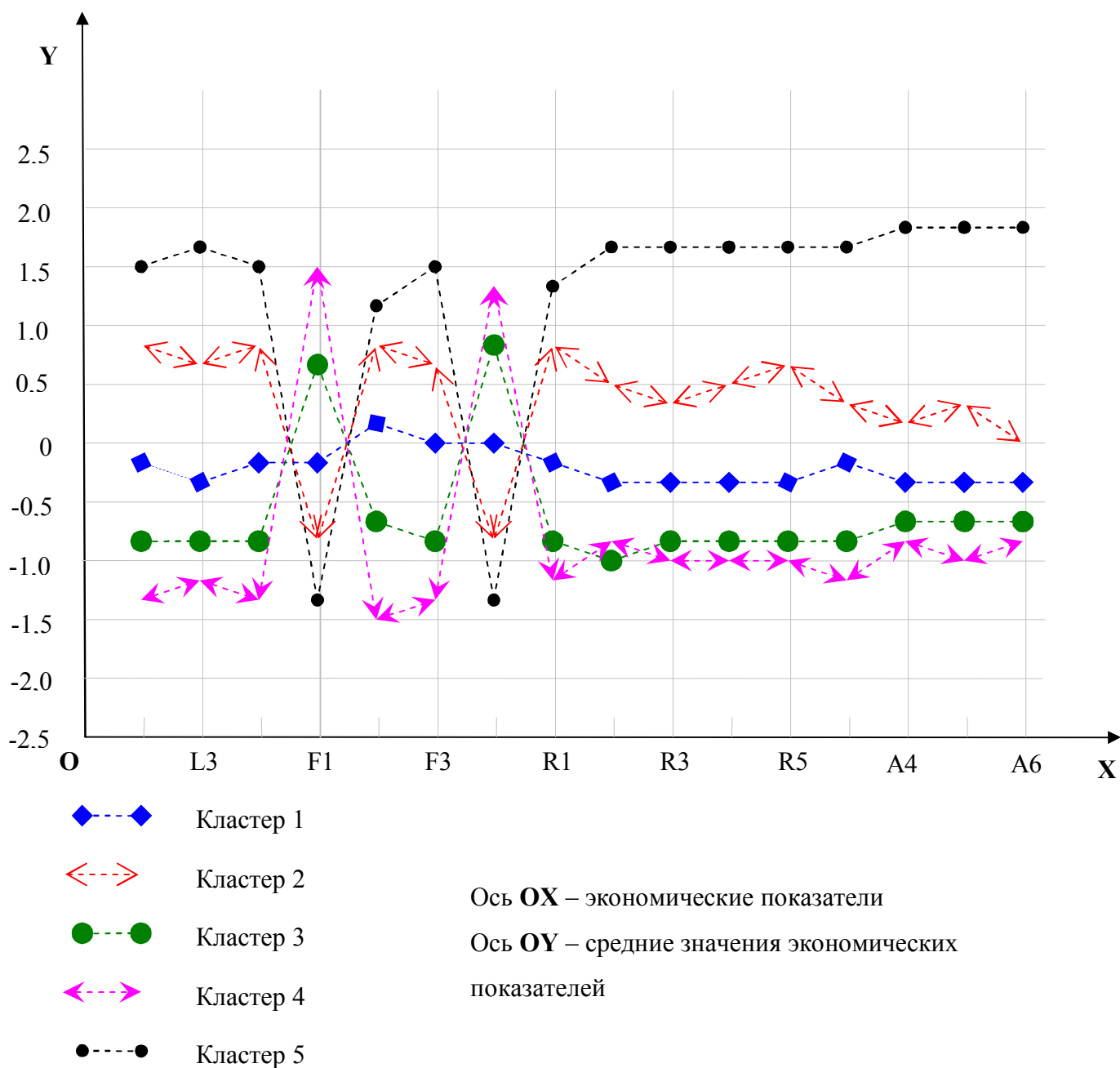


Рисунок 2. Графики средних значений показателей по кластерам

Из рисунка 2 видно, что средние всех показателей значительно отличаются друг от друга: расстояние между средними характеристик кластеров большое, и общее расстояние между центрами кластеров значительно. Это свидетельствует о качественном разбиении на группы при  $m=5$ , а значит и об успешной кластеризации.

Средства программы Statistica 6 позволили в файле сохраненных результатов кластерного анализа каждой из групп присвоить соответствующее значение лингвистической шкалы. Сравнение процесса генерации виртуальной базы предприятий с полученным разбиением на кластеры позволило убедиться, что была проведена 100%

успешная кластеризация: все предприятия попали в кластеры, соответствующие 5 классам уровней риска. В таблице 2 приведены примеры 10 предприятий из результирующего файла по 2 предприятия каждого уровня риска.

Таблица 2. Соответствие предприятий группам уровня риска

№ предприятия в общей базе	Группы уровня риска	Экономические показатели							
		L1	L3	P1	F1	F2	F3	F4	R1
1	Очень высокий риск	0,129	46,427	0,327	2,576	0,269	-1,821	1,567	0,422
2	Очень высокий риск	0,121	69,882	0,184	2,594	0,277	-1,846	1,561	0,350
201	Высокий риск	0,459	100,754	0,627	1,844	0,391	-1,304	1,360	3,579
202	Высокий риск	0,458	88,041	0,508	1,929	0,388	-0,968	1,329	3,573
401	Средний риск	0,977	161,788	1,335	1,328	0,679	0,484	1,022	13,354
402	Средний риск	0,961	168,287	1,311	1,573	0,693	0,492	0,968	11,667
601	Низкий риск	1,593	400,141	1,842	1,035	0,843	1,531	0,711	21,001
602	Низкий риск	1,570	376,547	2,040	1,087	0,849	1,216	0,607	22,461
801	Очень низкий риск	2,055	507,283	2,585	0,659	0,940	2,411	0,273	29,050
802	Очень низкий риск	2,100	527,547	2,473	0,565	0,943	2,721	0,377	28,846
№ предприятия в общей базе	Группы уровня риска	Экономические показатели							
		R2	R3	R4	R5	A2	A4	A5	A6
1	Очень высокий риск	-2,800	-3,533	0,451	-4,664	0,066	0,452	0,631	1,072
2	Очень высокий риск	-2,709	-3,848	0,369	-4,628	0,065	0,465	0,655	1,187
201	Высокий риск	-1,068	-1,462	2,264	-2,407	0,102	0,612	0,743	1,622
202	Высокий риск	-1,196	-2,344	3,345	-2,572	0,099	0,674	0,850	1,801
401	Средний риск	-0,053	1,423	8,362	0,368	0,172	1,261	1,464	2,711
402	Средний риск	0,000	1,288	8,736	0,315	0,169	1,213	1,461	2,673
601	Низкий риск	-0,529	5,052	24,470	8,166	0,264	1,860	1,892	4,852
602	Низкий риск	-0,484	6,523	21,198	5,400	0,278	1,697	2,087	4,807
801	Очень низкий риск	6,995	10,370	36,434	12,511	0,443	5,252	3,195	13,795

802	Очень низкий риск	6,092	12,274	35,281	14,397	0,390	5,130	4,180	10,467
-----	-------------------	-------	--------	--------	--------	-------	-------	-------	--------

В работе были изучены и учтены положительные и отрицательные стороны различных лингвистических оценок, что позволило выбрать наиболее оптимальный классификатор для решения поставленной нами задачи. Для выбранного классификатора была составлена таблица интервалов значений финансовых показателей предприятий. На основе данной таблицы – сгенерирована виртуальная клиентская база данных по 1000 предприятиям, которая позволит решить проблему информационной неопределенности в процессе определения уровня риска реально существующего предприятия. Произведя кластерный анализ базы, убедились, что сгенерированные данные качественно определяют принадлежность предприятий к обозначенным классам.

Данные исследования были проведены для макроэкономического периода с 2000 по IV квартал 2008 года, но они могут являться шаблоном для проведения аналогичных исследований любого другого макроэкономического периода. При анализе нового макроэкономического периода необходимо будет произвести новое качественное разбиение финансовых показателей на группы лингвистической шкалы согласно новой сложившейся макроэкономической обстановке в стране. Для этого потребуются повторный анализ показателей предприятий, аналогичный тому, что был уже произведен в работе [3]. Такой анализ возможен только по истечении времени, необходимого для сбора статистической информации о предприятиях.

### Список литературы

1. Недосекин А.О., Бессонов Д.Н., Лукашев А.В. «Сводный финансовый анализ российских предприятий за 2000-2003 г.г.»
2. Недосекин А.О. Применение теории нечетких множеств к задачам управления финансами // *Аудит и финансовый анализ*, №2, 2000.
3. Давыдова Г.В., Беликов А.Ю. Методика количественной оценки риска банкротства предприятий // *Управление риском*, 1999 г., № 3, с. 13-20
4. Сайт Международного валютного фонда [www.mvf.su](http://www.mvf.su).
5. Недосекин А.О. и Фролова С.Н. «Лингвистический анализ гистограмм экономических факторов»
6. Недосекин А.О. Финансовый анализ в условиях неопределенности: вероятности или нечеткие множества? // 1999.- На сайтах: [www.vmgrousp.ru](http://www.vmgrousp.ru), [cfin.ru/analysis](http://cfin.ru/analysis), <http://www.delovoy.newmail.ru/analitic/3.htm> .

7. А.А.Халафян STATISTICA 6. Статистический анализ данных. 3-е изд. Учебник – М.:ООО «Бином-Прогресс», 2007г. – 512 с.: ил.

### References

1. *Nedosecin A.O., Bessonov D.N., Lukashev A.V.* «Financial analysis of Russian enterprises»
2. *Nedosecin A.O.* Application of fuzzy sets approach to financial direct tasks. // *Audit and financial analysis*, № 2, 2000.
3. *Davidova G.V., Belicov A.U.* Quantitative valuing of insolvency risks methods for enterprises // *Risks direct*, 1999, № 3, pp 13-20.
4. Sait of International currency fond [www.mvf.su](http://www.mvf.su).
5. *Nedosecin A.O., Frolova S.N.* «Linguistic analysis of economic factor's gistograms»
6. *Nedosecin A.O.* Financial analysis of fuzzy conditions: probabilities or fuzzy sets?// 1999.
7. *Halaphyan A.A.* STATISTICA 6. Statistic data analysis. // 2007, pp 512